Received: 9 June 2020 | Accepted: 14 June 2020

DOI: 10.1111/imr.12897

INVITED REVIEW

Immunological Reviews WILEY

The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity

Inna G. Ovsyannikova | Iana H. Haralambieva | Stephen N. Crooke | Gregory A. Poland | Richard B. Kennedy

Mayo Clinic Vaccine Research Group, Mayo Clinic, Rochester, MN, USA

Correspondence
Richard B. Kennedy, PhD, Co-Director – Vaccine Research Group, Mayo Clinic, Guggenheim Building 611C, 200 First Street SW, Rochester, MN 55905, USA.
Email: kennedy.rick@mayo.edu

Funding information
National Institute of Allergy and Infectious Diseases, Grant Numbers: AI132348, AI121054, AI127365

Abstract

This article provides a review of studies evaluating the role of host (and viral) genetics (including variation in HLA genes) in the immune response to coronaviruses, as well as the clinical outcome of coronavirus-mediated disease. The initial sections focus on seasonal coronaviruses, SARS-CoV, and MERS-CoV. We then examine the state of the knowledge regarding genetic polymorphisms and SARS-CoV-2 and COVID-19. The article concludes by discussing research areas with current knowledge gaps and proposes several avenues for future scientific exploration in order to develop new insights into the immunology of SARS-CoV-2.

KEYWORDS

alleles, coronavirus, COVID-19, genes, genetic variation, genome-wide association study, GWAS, HLA, immunogenetics, polymorphisms, single nucleotide, SARS, SARS-CoV-2, severe acute respiratory syndrome, systems biology, vaccine

1 | INTRODUCTION

The 2019 outbreak of SARS-CoV-2, which began in Wuhan, China, refocused the world's attention on coronaviruses. Several seasonal coronaviruses are known to circulate in the human population and generally cause relatively mild respiratory tract infections, including the alphacoronaviruses HCoV-NL63 and HCoV-229E, and the betacoronaviruses HCoV-OC43 and HCoV-HKU1. Two of these viruses (OC43 and 229E) were first discovered in the late 1960s, while the other two (NL63 and HKU1) were identified in 2004 and 2005, respectively.[1] Phylogenetic evidence suggests that each of these four viruses originated from bat- (NL63 and 229E) or rodent-associated (OC43 and HKU1) coronaviruses.[1] These four seasonal coronaviruses account for 15%-30% of common colds in children and a smaller percentage of colds in adults. While seasonal coronavirus infections are generally mild, they can cause more severe disease in neonates and immunocompromised subjects.[2,3] It is established that adults have a >90% seroprevalence rate (for at least one of the seasonal coronaviruses), while children are susceptible and are exposed to infection during early childhood.[2]

In addition to these four seasonal coronaviruses, since 2003, three additional coronaviruses (ie, SARS-CoV-1: the severe acute respiratory syndrome coronavirus; MERS-CoV: the Middle East respiratory syndrome [MERS] coronavirus; and SARS-CoV-2: the causative agent of COVID-19) have emerged as human pathogens, and each is associated with severe infection.

SARS, caused by SARS-CoV-1, is a life-threatening lower respiratory tract infection (with atypical pneumonia and progressive lung damage) with significant morbidity and mortality that accounted for 8098 laboratory-confirmed cases and 774 deaths during 2002-2004.[4] SARS is believed to have arisen in bats, transferred to animals such as civet cats, and was then acquired by humans.[5] No further known cases have occurred since 2004.

MERS-CoV is a zoonotic betacoronavirus transmitted to humans from dromedary camels, with some evidence that it originated in bats.[6] MERS causes a wide range of infections in humans (from asymptomatic

This article is part of a series of reviews covering B and Th cell response to Ag in vivo: implications for vaccine development and diseases appearing in Volume 296 of *Immunological Reviews*.

to severe pneumonia with high mortality rate of 34%). It originated in Saudi Arabia in 2012 and has led to 2494 confirmed cases and 858 associated deaths.[7] Unlike SARS, which disappeared after 2004, this virus continues to be a problem, with sporadic small numbers of cases continuing to occur primarily within the Arabian Peninsula.

In 2019, the seventh coronavirus known to infect humans was identified as the cause of the pandemic originating in Wuhan, China. This virus, SARS-CoV-2, was initially associated with a respiratory disease termed COVID-19. Subsequently, a wide range of clinical outcomes have been observed among individuals diagnosed with COVID-19, ranging from mild respiratory infection to acute respiratory disease and death.[8-10] Disease severity is disproportionately higher among older adults and individuals with underlying comorbidities,[11] although severe cases of COVID-19 have also been reported among young and healthy individuals.[12] Symptomatology also varies widely, with case reports indicating gastrointestinal (GI) distress, sensorineural impairment, loss of smell and taste, dysregulated blood clotting, hepatitis, and acute cardiac and renal injury irrespective of age and health status.[13-19] The mechanisms underlying the variation in COVID-19 susceptibility and disease presentation are currently unknown, although viral and host genetic variants are probable factors influencing both disease severity and immune response outcomes (Figure 1).

Both candidate gene and genome-wide association studies (GWAS) have offered significant insights into the genetic basis of many infectious diseases. These studies identified genetic loci and allelic polymorphisms that determine, in part, genetic susceptibility to infections. Studies of viral and host genetics are critical for understanding the pathophysiology of SARS-CoV-2, elucidating why COVID-19 manifests differently among individuals, and informing the design of new vaccines and antiviral therapeutics. Such studies can help pinpoint if emerging viral strains are linked to more severe clinical outcomes or if individuals harboring certain alleles are more or less susceptible to disease. A seminal example is the loss-of-function mutation identified in the CCR5 gene (*CCR5Δ32*), which abrogates expression of CCR5 on the host cell surface and renders homozygous individuals resistant to infection from HIV.[20] Identification of such associations for SARS-CoV-2 could help guide the design of vaccines and therapeutics as well as prioritize individuals for treatment once an effective vaccine or antiviral therapy is developed.

The HLA region of the human genome has been recognized for its importance in both disease risk and resistance.[21] HLA gene polymorphisms have been linked to numerous infectious diseases, including those caused by RNA viruses, such as SARS, influenza, HIV, hepatitis C, rabies, West Nile fever, rubella, mumps, measles, and

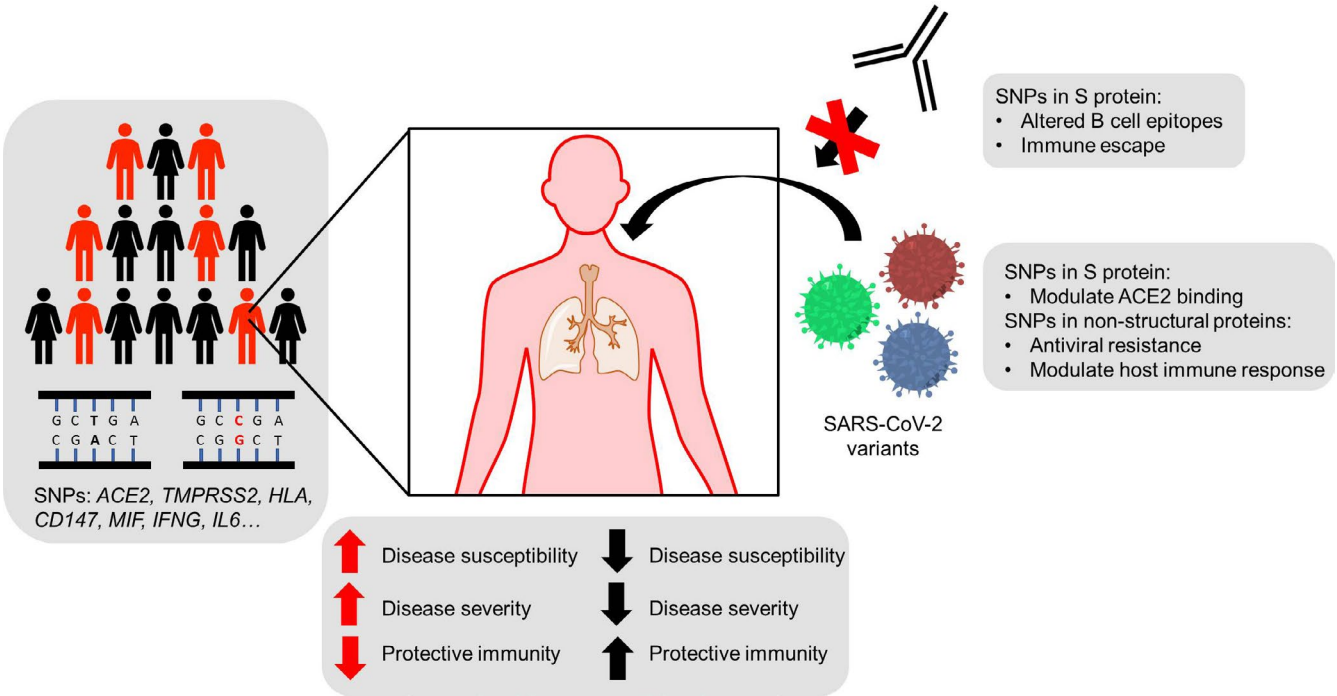

FIGURE 1 The impact of host genetics and viral variation on SARS-CoV-2 infection and COVID-19 severity. Individuals in the population harbor single nucleotide polymorphisms (SNPs) across a variety of genes (eg, *ACE2*, *TMPRSS2*, *HLA*, *CD147*, *MIF*, *IFNG*, *IL6*) that have been implicated in the pathology and immunology of SARS-CoV-2 and other pathogenic coronaviruses. These and other genetic variants may modulate disease susceptibility, increase or decrease disease severity, alter the variety of symptoms developed, and affect the magnitude and/or quality of the immune responses against SARS-CoV-2. In addition to host genetic variation, genetic variants of SARS-CoV-2 (and other pathogenic coronaviruses) can exhibit differences in biological activity. Single amino acid mutations in the spike glycoprotein can modulate ACE2 binding or alter B cell epitopes to promote immune escape or render monoclonal antibodies ineffective, while mutations in non-structural/accessory proteins can promote the development of resistance to antivirals, alter T cell epitopes, disrupt cell mediated immunity, and modulate host cellular interactions with viral particles

others. Such genetic association studies and the insights they provide are critical to identifying the HLA alleles that may be associated with protective immune responses. Differential immune responses have been observed in mild and severe cases of COVID-19, including delayed IgM responses[22] and higher S protein IgG titers in non-ICU patients,[23] raising the questions of if and how HLA allelic variation contributes in this differential immunity. It is critical to determine and compare HLA profiles among individuals with serious cases of COVID-19 with those who have mild or no apparent disease. This will aid in understanding the underlying mechanisms of protective innate and adaptive immunity and may lead to the development of genetic markers as correlates of protection.

Herein, we review the current status of research focused on identifying genetic variants associated with the immune response and clinical outcomes of coronaviruses, with particular attention to the contemporary SARS-CoV-2 pandemic. Limited data are available for the seasonal coronaviruses and for MERS-CoV, with more known for SARS-CoV-1. Specific to SARS-CoV-2, we are still in the very early stages of the COVID-19 pandemic, reports in the literature are largely preliminary, and studies have largely been conducted with computational methodologies and relatively small sample sizes due to the limited availability of biological samples. Nonetheless, several important findings have been reported for both viral and host genetic variants, and large consortiums (COVID-19 Host Genetics Initiative, COVID Human Genetic Effort) are being organized to further study the influence of genetics on SARS-CoV-2 infection and immune response. After reviewing the extant literature on host genetics and coronaviruses, we discuss currently unanswered questions and outline a proposed research agenda that will address those knowledge gaps.

2 | HOST GENETIC FACTORS AND SEASONAL CORONAVIRUS SUSCEPTIBILITY, PATHOGENESIS, AND HOST IMMUNE RESPONSE

From the perspective of basic science, previous studies have elucidated host factors that impact coronavirus host-pathogen interactions, coronavirus entry, replication, and pathogenesis, as well as modulation of cell cycle and host response (ie, innate and inflammatory response), as nicely summarized by de Wilde et al.[4] However, only a few studies have examined host genetic variation in the human population to answer critical questions about observed differences in susceptibility, clinical course, and outcome of coronavirus infections in humans. Previous candidate gene association studies have provided valuable information regarding what role human genetic factors play in determining how different variables influence coronavirus infections (eg, genetic predisposition, pathogenesis, clinical course, outcome), with the focus predominantly on SARS. Large, systematic studies assessing genome-wide associations of genetic variants with coronavirus disease in humans are lacking. An interesting genome-wide assessment of host genetic loci associated with SARS infection in mice (published in PLoS Genetics by Gralinski et al) was carried out using the Collaborative Cross mouse panel approach, allowing expanded opportunities for evaluation of SARS-related features (eg, lung pathology, viral titer, weight loss) in mice susceptible or resistant to coronavirus infection.[24] Genetic mapping in this study revealed several loci of interest, including a locus on chromosome 3 (with 23 genes and 13 non-coding RNAs) that contained the ubiquitin E3 ligase *TRIM55* gene (tripartite motif containing 55), a RING zinc finger-containing protein with possible engagement in protein-protein interactions, and a plausible role in SARS-CoV-induced vascular cufflink/inflammatory response in the lungs.[24] The implications of these findings for human coronavirus disease are unclear, and the impact of this locus/gene on coronavirus infection in humans has yet to be determined.

While much is known about the epidemiology, surveillance, transmission, shedding, and clinical manifestations of infections caused by the community-acquired viruses HCoV-NL63, HCoV-229E, HCoV-OC43, and HCoV-HKU1, limited data exist in the literature to elucidate the possible human genetic variation that may impact common cold coronavirus susceptibility and disease.[2,25-27] One common target for studies evaluating host genetic factors is the viral receptors. Early animal studies in a murine model of mouse hepatitis coronavirus (MHV) provided proof-of-concept evidence for the critical role of the coronavirus entry receptor genotype/alleles (ie, *CEACAM1*) for susceptibility to coronavirus infection.[28,29] Aminopeptidase N (*CD13*) is a cell surface metalloprotease that has been previously recognized as a receptor for HCoV-229E,[30] while ACE2 has been identified as a receptor for HCoV-NL63[31,32]; however, genetic variation in these receptors has not been studied in relation to these infections.

Recent knowledge of the mechanisms and host factors influencing infection by seasonal and zoonotic coronaviruses has been reviewed by Wilde et al.[4] The authors summarize the important role of host factors associated with coronavirus cell entry and fusion (ie, coronavirus receptors and cell surface or lysosomal proteases: CEACAM1, ACE2, APN, DPP4, TMPRSS2, cathepsins, and furin), host antiviral/proinflammatory, translation, and unfolded protein response factors (eg, eIF4F, GCN2, PERK, PKR, RIG-I, MDA-5, TLR3, IRF3, IRF7, type I interferon and proinflammatory cytokines, and chemokines), and host factors influencing coronavirus replication/protein expression (eg, MADP1, ANXA2, cyclophilins) for regulating coronavirus infection.[4] Of the listed factors, cyclophilins are of particular interest because their genetic variability has been associated with coronavirus infection in humans.[33] Cyclophilins are peptidyl-prolyl isomerases involved in folding of cellular proteins and coronavirus proteins and are necessary for proper viral propagation. It has been demonstrated that cyclophilin A directly interacts with the nucleocapsid of SARS-CoV-1[34] and is incorporated into the purified virions.[35] Functional single nucleotide polymorphisms in cyclophilin A may directly affect the propagation of HCoV-229E.[33] Data on the host factors associated with the seasonal coronavirus infections may provide valuable insights into the genetic factors associated with SARS and COVID-19 susceptibility and progression.

3 | HOST GENETIC FACTORS ASSOCIATED WITH SARS-COV-1 SUSCEPTIBILITY, PATHOGENESIS, AND HOST IMMUNE RESPONSE

Previous findings with SARS-CoV-1 have offered important—although often conflicting—insights regarding the plausible genetic influence of host factors on the susceptibility, pathogenicity, and clinical outcome of SARS. Most of these studies, which encompassed cohorts with limited sample size, relied on candidate gene approaches to investigate the genetic contribution of factors with a known or suspected role in coronavirus infection and pathogenesis. Because the angiotensin-converting enzyme 2 (*ACE2* mapping on the human chromosome Xp22, and a homologue of ACE1 with 40% amino acid identity) has been established as an entry receptor for at least three coronaviruses (ie, SARS-CoV-1, HCoV-NL63, and SARS-CoV-2[36-41]), several studies have explored the impact of *ACE2* polymorphisms on SARS susceptibility and disease severity.[42,43] A candidate gene case-control study by Chiu et al,[42] which explored the link between common *ACE2* single nucleotide polymorphisms/SNPs and SARS in a cohort of 168 SARS patients and 328 healthy controls of Chinese ethnicity, found no evidence for associations of *ACE2* genetic variants with SARS susceptibility, clinical manifestations, or clinical outcome. In a study from Vietnam, 44 SARS cases, 16 antibody-positive contacts, and 137 other controls were investigated for the genetic association between 19 SNPs in or flanking the *ACE2* gene and found no evidence for genetic association.[43] A recent GWAS study used a cohort of limited sample size (HCV-infected liver tissue samples from 195 subjects) to investigate associations between host genetic polymorphisms and *ACE2* gene expression.[44] The study found that a locus of genetic variation on chromosome 19 that controls the expression of *IFNL3* and *IFNL4* is also associated with *ACE2* expression, as was age.[44] These findings suggest the negative correlation between interferon response and *ACE2* expression, which may influence viral entry and infection by viruses using the ACE2 receptor.

The role of the type II transmembrane protease TMPRSS2 and other host proteases involved in SARS Spike (S) protein cleavage and activation to promote efficient infection[45,46] has not been studied in terms of host genetic heterogeneity. One highly cited SARS genetic risk assessment study[47] investigated the role of a specific *CLEC4M* gene polymorphism (in the variable tandem repeats in exon 4) in affecting the susceptibility and severity of SARS, assuming the encoded protein L-SIGN mediates or facilitates virus attachment and entry.[48] The study encompassed 285 confirmed SARS cases from Hong Kong and three cohorts of controls: 380 random healthy blood donors; 290 SARS-negative patients from outpatient clinics; and 172 SARS-negative healthcare workers. The results provided evidence for the protective role of the *CLEC4M* tandem repeats polymorphism against SARS. The C-type lectin domain family 4 member M (*CLEC4M*) (L-SIGN/CD209L) gene encodes a protein, which is highly expressed by endothelial cells in lymph nodes and liver, that can serve as an adhesion molecule/receptor for viruses (eg, human immunodeficiency virus [HIV], hepatitis C, Ebola virus [EBOV], and SARS-CoV-1) by binding carbohydrate ligands such as high-mannose oligosaccharides.[47-52] The function of this protein as a pathogen recognition receptor with direct involvement in SARS coronavirus infection makes its role as a host genetic restriction factor highly plausible. Chan et al[47] provided functional evidence for this hypothesis, demonstrating the differential impact of L-SIGN homozygosity for SARS-CoV adhesion, infection, proteasome-related viral degradation, and modulation of viral replication/titers.[47] The enthusiasm for these findings, however, was dampened by further studies that failed to replicate these results.[53,54] A genetic association study of the *CLEC4M* tandem repeats polymorphism[53] included case-control samples from northern China (a total of 441 SARS cases and 396 controls) and did not find a significant association between *CLEC4M* genotypes, homozygote or heterozygote frequencies, and SARS. Similarly, a study investigating the genetic predisposition for SARS with a focus on the C-type lectin cluster at chromosome 19p13.3 (*FCER2*, *CLEC4G*, *CD209*, *CLEC4M*) found no evidence of an association.[54] Several candidate genetic association studies have explored the association between genetic variants in innate immune response and other immune function-related genes, including antiviral effectors and chemokines/cytokines (*MBL*, *FCGR2A*, *MX1*, *OAS1*, *IL12RB1*, *IFNG*, *CCL2*, *CCL5/RANTES*, *DC-SIGN/CD209*, *ICAM3*) and susceptibility, pathogenesis, and disease course of SARS. It should be noted that many of these have not been replicated, and some studies show conflicting results.[55-64] Of interest are several studies investigating the link between SARS and mannose-binding lectin (MBL) deficiency, as well as the impact of genetic variants of the mannose-binding lectin/mannan-binding lectin gene (*MBL2*), encoding a pattern recognition innate protein/collectin that is instrumental in the inactivation of a variety of respiratory pathogens through direct binding (to repeated mannose and N-acetylglucosamine entities) and complement activation.[55,57,65,66] Human clinical studies have provided evidence that MBL deficiency and *MBL2* polymorphisms are associated with morbidity and death as a result of respiratory and other severe infections such as pneumococcal pneumonia, tuberculosis, and meningococcal disease.[65,67-70] A large case-control study by Ip et al,[55] which included 569 SARS patients and 1188 controls, demonstrated that lower serum levels of MBL and MBL deficiency are host factors associated with increased susceptibility to SARS. It was found that the median serum MBL in SARS patients was 0.733 μg/mL, which is significantly lower than the MBL level found in healthy control subjects (1.369 μg/mL, *P*-value = .0004). The impact of the *MBL* X/Y promoter polymorphisms and the structural A/B polymorphisms, as well as the three *MBL* haplotypes (YA, XA, and YB) on SARS susceptibility, was evaluated in this study. The haplotype YB, associated with MBL deficiency, was found to be more frequent in SARS patients (33.4%) compared to 24.7% frequency in the control group (*P*-value < .001, odds ratio/OR 1.52).[55] The study also elucidated some of the mechanisms underlying MBL effects, such as direct binding of MBL to

SARS-CoV-1, complement activation, and MBL-mediated inhibition of viral infection.[55] A similar study by Zhang et al[57] studied 353 SARS patients and 392 controls to investigate the genetic influence of several promoter *MBL2* polymorphisms and the codon 54 (exon 1) *MBL2* polymorphism (rs1800450) on SARS susceptibility. The results provided evidence for the significant association (*P*-value = .00085) between the codon 54 *MBL2* polymorphism (resulting in diminished expression of MBL, *P*-value = .00187) and predisposition/susceptibility to SARS.[57] However, other studies have found no significant difference in MBL genotypes/haplotypes between SARS patients (n = 180 SARS patients from Hong Kong, including 132 patients with moderate disease, 26 patients with severe disease, and 22 patients with lethal SARS) and 200 control subjects.[56] This study reported a possible link between SARS severity and outcome and discovered a non-synonymous polymorphism affecting amino acid position 131 (arginine or histidine residue) of the Fc fragment of IgG receptor IIa (*FCGR2A* gene). The FcgammaRIIA-receptor R131 genotype was found to be significantly associated with SARS severity (*P*-value = .03), with a higher homozygosity frequency of this genotype among the ICU SARS patients compared to controls.[56] The encoded receptor is found on macrophages/neutrophils and other cells and is directly involved in innate/inflammatory response processes such as phagocytosis.[56] A study from Hong Kong included SARS patients (n = 495) and healthy controls (n = 578) who were genotyped for a panel of SNP markers, such as *IFNγ*, *IL-10*, *TNFα*, *IL-12*, *RANTES*, *IL-10*, *Mig*, and *MCP-1* genes.[71] Polymorphisms in the *IFNγ + 874A* and *RANTES-28G* alleles have been found to be significantly associated with SARS susceptibility and death in patients with SARS.[71] The IFNG + 874A allele was found more frequently in SARS patients (83.1%) compared to controls (66.3%, *P*-value < .001). This allele was significantly associated with SARS susceptibility (ie, subjects with IFNG + 874 AA or AT genotype had the OR of 5.19 and 2.57, respectively, of developing SARS [*P*-value < .001]).[71] In addition, subjects with the RANTES −28 GC and GG genotypes had the OR of 3.28 and 3.06, respectively, of developing SARS (*P*-value < .0001).[71] High circulating levels of the cytokine macrophage migration inhibitory factor (MIF), which is expressed on alveolar epithelial, endothelial, and many immune cells, have been implicated in the pathogenesis of influenza, West Nile, and HIV infections.[72] For example, functional polymorphisms in the *MIF* gene (CATT7 and −173 C alleles) have been previously reported as potential predictors of community-acquired pneumonia and morbidity/mortality outcomes in patients with pneumococcal meningitis.[73,74] Given the morbidity and mortality of COVID-19, it seems important to investigate the role of *MIF* gene polymorphisms in disease severity and clinical outcomes.

While the listed studies provide valuable insight into plausible host genetic factors influencing SARS susceptibility, pathogenesis, and outcome, no systematic GWAS (agnostic to previous knowledge) for coronavirus infections in humans has been published.

The influence of human leukocyte antigen (HLA) gene polymorphisms for SARS susceptibility, pathogenesis, and outcome has also been investigated in a number of studies, predominantly in Asian populations. Several HLA class I polymorphisms (ie, HLA-B*46:01,[75] HLA-B*07:03,[76] and HLA-Cw*08:01[77]) have been significantly associated with susceptibility to SARS and/or disease severity in various populations (HLA-B*46:01 *P*-value = .04 for association with SARS susceptibility and *P*-value = .008 for association with SARS severity[75]; HLA-B*07:03 *P*-value = .00072 for association with SARS susceptibility[76]; and HLA-Cw*08:01 *P*-value = .007 for association with SARS susceptibility[77]). Of these, the HLA-B*46:01 allele was computationally predicted to bind to the fewest SARS-CoV-2 peptides, thus suggesting it as a non-protective allele for COVID-19 disease.[78] Similarly, HLA class II polymorphisms (HLA-DRB4*01[79] and HLA-DRB1*12:02[80]) were demonstrated to be significantly associated with the predisposition for SARS infection (HLA-DRB4*01 *P*-value = .0031 for association with SARS susceptibility[79]; HLA-DRB1*12:02 *P*-value = .0065 for association with SARS infection[80]). The protective effect of HLA-DRB1*03:01,[81,82] HLA-Cw*15:02,[81] and HLA-A*02:01[82] has been suggested (*P*-value < .05) by studying susceptibility/resistance to SARS in cases-control studies. Other studies, however, found no evidence of association between SARS susceptibility/disease and HLA-A, HLA-B, and HLA-DRB1 loci in subjects of Chinese ethnicity.[83,84] Comprehensive, large, and reproducible studies are still needed to decipher any possible genetic predisposition underlying susceptibility to SARS and disease progression/host immune response.

4 | HOST GENETIC FACTORS ASSOCIATED WITH MERS SUSCEPTIBILITY, PATHOGENESIS, AND HOST IMMUNE RESPONSE

With MERS, a variety of host factors associated with disease susceptibility and virus transmission have been identified, including the virus entry receptor (dipeptidyl peptidase-4 [*DPP4*]), presumed attachment factors, sialic acids, host proteases (eg, TMPRSS2, furin, cathepsins), interferons, interferon-stimulated genes, and adaptive immune response factors.[85-87] Unlike for SARS, there are few published reports evaluating host genetic variations associated with MERS susceptibility, pathogenesis, transmission, and morbidity/mortality, with only one published study in the literature.[88] This study, which examined HLA class II alleles in 23 MERS patients and 161 healthy subjects from Saudi Arabia, concluded that HLA-DRB1*11:01 (*P*-value = .0016) and HLA-DQB1*02:02 (*P*-value = .027) alleles were associated with susceptibility to MERS, but not with disease pathogenesis and outcome.[88] More comprehensive genetic association studies with larger sample sizes are warranted to both validate these preliminary findings and to further investigate the contribution of HLA to MERS disease susceptibility, clinical course, and sequelae, as well as to assess the role of genetic variation in other crucial host factors (eg, *DPP4*, *TMPRSS2*).

In summary, the immunogenetics of coronavirus infections are largely understudied, and high-quality genetic association studies, genetic variant functional studies, and systems biology studies have not been performed, but they are needed in order to provide novel and important data regarding the genetic control of coronavirus infection, as well as host response and immunity.

5 | HOST GENETIC FACTORS AND SARS-COV-2 SUSCEPTIBILITY, PATHOGENESIS, AND HOST IMMUNE RESPONSE

As has been described for the other coronaviruses (seasonal, SARS, and MERS), host genetic variation may be a key factor influencing the susceptibility, severity, and overall clinical outcomes of COVID-19 after infection with SARS-CoV-2. It is well known that individuals of diverse racial and ethnic backgrounds harbor different allelic variants,[89-91] and gene expression differs based on the biological age and sex of an individual.[92-94] In the few short months since the beginning of the COVID-19 pandemic, researchers have already reported several lines of evidence suggesting genetic factors impact COVID-19 severity. These include early reports suggesting that males were more likely to suffer from severe disease,[10,95-97] reports that blood type A was a risk factor for severe disease,[98,99] and clear racial differences in clinical outcomes observed in the United States.[100,101] It has also been suggested that the reduced severity of disease in females may be due to antibody responses that develop more quickly and with higher titers than antibody responses in men following infection.[102]

Past reports from studies of other coronaviruses have highlighted the important role that cellular receptors play in disease. Individuals carrying specific variants of genes directly involved in viral infection (eg, *ACE2*, *TMPRSS2*) or exhibiting differential expression of those genes may have inherently different susceptibility to SARS-CoV-2, which may explain the broad spectrum of symptoms and disease severity associated with COVID-19.

5.1 | ACE2

As ACE2 is the primary cellular receptor for SARS-CoV-2, studies have begun to investigate the relationship between *ACE2* polymorphisms and COVID-19 severity. A recent study by Renieri and colleagues integrated whole-exome sequencing data with molecular dynamics simulations to identify and characterize 33 ACE2 variants from ~7000 Italian subjects.[103] Notably, one variant (N720D) was proximal to the TMPRSS2 cleavage site, while three mutations (W69C, L351V, P389H) were predicted to induce conformational changes that alter interactions with the RBD of the S glycoprotein. Stawiski et al also curated a list of *ACE2* polymorphisms from existing genomic datasets and used structural modeling to identify variants that putatively increase (S19P, I21V, E23K, K26R, T27A, N64K, T92I, Q102P, H378R) or decrease (K31R, N33I, H34R, E35K, E37K, D38V, Y50F, N51S, M62V, K68E, F72V, Y83H, G326E, G352V, D355N, Q388L, D509Y) COVID-19 susceptibility on the basis of interactions with the S glycoprotein.[104] A recent study of European and East Asian cohorts also identified two *ACE2* variants (K26R and I468V) as having potentially lower binding affinities for the S protein.[105] Based on their proximity to important interacting regions on the S protein surface, the variants identified in these studies have the potential to alter the kinetics of viral binding and internalization and may explain the significant observed inter-individual variability in COVID-19 case severity.

A study by Pinto and colleagues analyzed 700 lung transcriptomic datasets from patients with diabetes, hypertension, and chronic obstructive pulmonary disease and found that global *ACE2* expression was significantly higher than in the lungs of healthy patients.[106] This may explain why these individuals are predisposed to a higher severity of COVID-19; however, other studies suggest that higher ACE2 expression may be protective against lung injury, and ACE2 expression decreases with age (at least in an animal model).[107-109] Additional studies are required to determine if the modulation of ACE2 is a direct result of underlying comorbidities or if the increased/decreased expression of ACE2 in these individuals in certain anatomical locations/tissues is causal for severe COVID-19.

The increased severity of COVID-19 in males has also been speculatively associated with *ACE2* polymorphisms and expression levels. Expression of *ACE2* has generally been found to be higher in men than women,[110,111] although some studies have reported equivocal expression levels.[112,113] These conflicting data have led some studies to suggest that it is either the pattern or density of *ACE2* expression in different anatomical locales rather than global differences in expression levels that influence sex-based differences in disease severity.[114-116] As *ACE2* is encoded on the X chromosome, males only carry and express a single ACE2 variant on all cells, whereas females will likely express a mosaic pattern of ACE2 as determined by early X-inactivation events.[114] Thus, if men harbor an *ACE2* variant that is more permissive for SARS-CoV-2 infection, all of their cells will express this risk variant.

5.2 | TMPRSS2

The serine protease TMPRSS2 is the second host protein involved in SARS-CoV-2 infection and has received considerably less attention in genetic association studies to date. Russo et al identified an intergenic single nucleotide polymorphism (SNP) that was concomitantly associated with increased expression of *TMPRSS2* and reduced expression of the interferon-inducible gene *MX1* in lung tissue.[117] Individuals carrying this SNP may have an increased susceptibility to SARS-CoV-2 infection due to the increased expression of TMPRSS2 on the cell surface and the simultaneous dampening of the cellular antiviral response. A second study by Asselta and colleagues identified numerous SNPs in *TMPRSS2* in an Italian cohort that were all predictively associated with higher gene expression levels.[112] Notably, one of these variants was associated with an

androgen-responsive enhancer upstream of *TMPRSS2*, suggesting another possible mechanism for the enhanced severity of COVID-19 in males. While the preliminary identification of both ACE2 and TMPRSS2 variants potentially associated with COVID-19 susceptibility is promising, studies investigating the associations between these genetic variants and clinical outcomes in individuals infected with SARS-CoV-2 are lacking. Thus, classification of identified variants as protective or deleterious is premature until additional studies validating their biology can be undertaken.

5.3 | HLA loci

Currently in the literature, there are limited reports of studies investigating HLA genetic variation and the immune response against SARS-CoV-2. A single study reported that HLA-A*24:02 was associated with COVID-19 susceptibility after identifying this allele in four of five patients from Wuhan,[118] although it is highly unlikely this is a reproducible finding given the small sample size. While biological data are limited, several studies have reported the use of predictive algorithms to identify HLA alleles associated with viral peptide epitope recognition.[78,119,120] These studies may provide some evidence as to which HLA alleles play important roles in mediating protection from SARS-CoV-2. Nguyen et al sampled peptides from the entire SARS-CoV-2 proteome across ~150 HLA class I genotypes to map susceptibility loci for COVID-19.[78] They identified HLA-B*15:03 as having a high capacity for presenting peptides conserved between SARS-CoV-2 and other pathogenic coronaviruses, suggesting that this allele may be broadly protective. Conversely, they predicted HLA-B*46:01 to bind to the fewest number of peptides from SARS-CoV-2, suggesting that individuals who carry this allele may mount a weaker immune response and develop more severe symptoms. This allele was previously predicted to be a susceptibility marker for SARS-CoV infection.[75] Specifically, in SARS coronavirus infection, the HLA-B*46:01 (*P*-value = .0008) has been significantly associated with the severity of SARS in Asian populations.[75] Large-scale GWAS and biological validation studies with larger cohorts of convalescent individuals are warranted to fully understand the immune response to SARS-CoV-2. In fact, a recent genome-wide study involving 1980 COVID-19 patients in Spain and Italy identified two genetic regions associated with SARS-CoV-2-induced respiratory failure.[121] One signal was located at 9q34 within the ABO blood group locus, with A-positive individuals having a higher risk (OR = 1.45, $P = 1.48 \times 10^{-4}$) and blood group O individuals at lower risk (O = 0.65, $P = 1.06 \times 10^{-5}$), confirming earlier reports linking blood type A to more severe disease and blood type O to a protective effect.[98,99] The other locus was found at 3p21.31 in a region including six genes (*SLC6A20*, *LZTFL1*, *FYCO1*, *CXCR6*, *XCR1*, *CCR9*), including chemokine receptors and the SLC6A20 gene that encodes for a transporter protein known to interact with ACE2. Following up on these findings may reveal additional insights into the influence of genetic factors on clinical outcomes to COVID-19.

5.4 | Inflammatory factors and IL-6

The critical role of exaggerated inflammatory responses leading to pathology associated with SARS-CoV-2, SARS-CoV-1, and MERS infections has been recently reviewed.[122] The release of proinflammatory cytokines (eg, IL-6, IL-1β, TNFα) in the course of infection, referred as "cytokine storm," is associated with the development of severe alveolar damage and lung inflammation/pathology characteristic of the acute respiratory distress syndrome.[123] Among the factors associated with increased COVID-19 severity and lethal outcome (ie, older age, comorbidities, lymphopenia, secondary infections), increased levels of ferritin, D-dimer, C-reactive protein, and increased cytokine levels of IL-2R, IL-6, IL-10, and TNF-α (*P*-value < .001; *P*-value = .04; *P*-value = .001 and *P*-value = .04, respectively, when comparing severe vs mild COVID-19 cases) were reported to correlate with COVID-19 disease severity.[124] In addition, chronic inflammation is one of the key characteristics of aging, obesity, and some of the associated comorbidities (eg, hypertension, diabetes) that influence the clinical course and outcome of COVID-19. The combination of antiviral therapeutics and cytokines is already in clinical trials for the treatment of mild-to-moderate COVID-19.[125] IL-6 is a central proinflammatory cytokine with pleiotropic functions and specific effects in the lung microenvironment during viral infections. It can have opposing effects by stimulating adaptive immune responses by antigen-specific B cells and CD8 + T cells and facilitating the survival of phagocytes, but it can also promote imbalanced Th2 and Th17 T cell differentiation over Th1 T cell differentiation, enhance lung tissue damage, edema, and vascular permeability, and facilitate the infiltration of proinflammatory macrophages and neutrophils.[122] It has been demonstrated that COVID-19 pneumonia is characterized by immune system dysregulation with excessive production of IL-6 and that elevated IL-6 levels were associated with and can predict respiratory failure and death in COVID-19 patients.[124,126-128] In view of the role of IL-6 in COVID-19-associated pathology, different anti-IL-6 therapeutics have been considered for the treatment of severe COVID-19, including a monoclonal antibody/Sarilumab, which blocks the IL-6 receptor, that is currently in clinical trials.[122] The first studies suggesting the role of IL-6 polymorphisms in the progression of COVID-19 are underway.[129,130] Using meta-analysis, a recent study analyzed 671 severe cases of bacterial and viral non-COVID-19 pneumonia and 2910 non-severe case of bacterial and viral pneumonia and found a significant association of the IL-6-174C allele (associated with a higher IL-6 level) with pneumonia severity (C allele vs G allele, OR: 1.33, 95% CI1.04-1.69, *P*-value = .019).[129] Given the importance of host genetic factors for infectious disease susceptibility and pathogenesis, the probable link between IL-6 genetic polymorphisms and COVID-19 warrants comprehensive investigation in large population based studies.

5.5 | COVID-19 host genetics initiative

While reports on the association between viral/host genetics and COVID-19 susceptibility are still scarce, massive initiatives are

currently being undertaken to promote research in this area. The COVID-19 Host Genetics Initiative[131] was recently established as a global resource for the generation and exchange of genomic data related to COVID-19 susceptibility.[132] As of this writing, more than 190 studies—including several large biobank repositories—are registered with this initiative. The COVID Human Genetic Effort,[133] another international consortium recently established by the National Institutes of Health and Rockefeller University, which is focused on identifying monogenic variations that underlie COVID-19 severity or resistance to SARS-CoV-2 infection. The Genetics of Mortality in Critical Care (GenOMICC) consortium[134] recently partnered with Genomics England, Illumina, and United Kingdom National Health Service to sequence the genomes of 20 000 patients hospitalized with COVID-19. Consumer genetics companies—such as 23andMe and Ancestry—have also initiated efforts to conduct GWAS studies related to COVID-19 susceptibility and disease severity. Large-scale studies and consortiums such as these will allow the rapid advancement of science in order to identify and better understand the genetic determinants of SARS-CoV-2 infection and immune response.

6 | SARS-COV-2 GENETIC VARIANTS

This review has focused on the effect of host genetic variation on disease susceptibility. Alterations in the virus genome, through mutation or recombination events, also have the potential to affect all aspects of the viral life cycle, including transmissibility, cellular tropism, and disease severity. In fact, mutations within the SARS-CoV-2 genome are the simplest explanation for the wide variety of clinical outcomes observed with COVID-19. The genomes of single-stranded RNA viruses accumulate mutations at a rate of 10^{-6}-10^{-4} per replication cycle,[135] which is significantly faster than the mutation rate for the human genome (10^{-8} per generation).[136] This leads to the accrual of numerous quasi-species within a single infected individual, which may account for the observed differences in symptoms and disease severity.[137] Mutations among viral progeny may result in altered ACE2 binding interactions or shifted tissue tropism that promotes more or less aggressive and widespread infections. Research is needed to investigate SARS-CoV-2 mutations within infected individuals to determine if there is a causal relationship between the emergence of specific mutations and disease severity or the spread of infection to specific tissues.

Certain mutations may provide a virus with distinct evolutionary advantages, such as changing a primary epitope to mediate escape from the host immune system or altering virulence factors to enhance transmission. These mutations may become established as a result of natural selection or vaccine selective pressure and subsequently give rise to new viral strains.[138] Forster et al reported the identification of three major variant types (A, B, C) of SARS-CoV-2 following phylogenetic analysis of 160 viral genomes.[139] Interestingly, the B-type viruses (T8782C + C28144T) were constrained to East Asia, while the A (T29095C) and C-type (T8782C + C28144T + G26144T) viruses were predominant in Europe and North America, which suggests that selective events have already occurred to allow the spread of SARS-CoV-2 variants beyond Asia. Real-time monitoring of global viral spread[140,141] suggests that as many as eight strains of SARS-CoV-2 are currently circulating, although they continue to maintain a high degree of sequence similarity. This is consistent with observations from epidemiologic studies in Italy and the United States, suggesting that only 4-10 stable non-synonymous mutations have been introduced into the SARS-CoV-2 genome since its emergence in Wuhan several months ago.[142,143] It is important to note that one such mutation in the S protein (D614G) appears to significantly increase the transmissibility of SARS-CoV-2, with strains harboring this mutation spreading rapidly throughout Europe and the United States since their origin.[144] It has been reported that individuals infected with the 614D variant have a higher case fatality rate that individuals infected with the 614G variant.[145] The authors suggest that differential S protein stability (and binding to ACE2) may play a role in the varying disease courses.

While the genome of SARS-CoV-2 appears to be relatively stable, continued surveillance for mutations remains crucial. Vaccine and therapeutic development could be undermined by the introduction of stable mutations that alter epitopes or antiviral binding sites, and mutations that alter infectious processes may give rise to strains that cause more (or less) severe disease. A limited number of studies have undertaken comparative analyses to monitor the emergence of new SARS-CoV-2 variants. Pachetti and colleagues characterized eight recurrent mutations across 220 clinical samples, one of which (P323L) was proximal to a putative antiviral binding site in the RNA-dependent RNA polymerase (RdRp).[146] Similar mutations in the RdRp of other viruses have been associated with resistance to antiviral therapy.[147,148] Jia et al recently reported phylogenetic analysis of the viral S glycoprotein from 106 SARS-CoV-2 isolates from 11 countries.[149] While the S glycoprotein was largely conserved among these samples, a variant from India was identified with a mutation (R408I) in the receptor-binding domain (RBD) and was predicted to significantly alter ACE2 binding affinity.[149] A larger study by Ou et al identified 32 non-synonymous mutations in the S glycoprotein RBD across 1609 viral genomes and employed molecular dynamics simulations to characterize their binding with ACE2.[150] Three of these variants (V367F, W436R, and D364Y) were predicted to have binding affinities for ACE2 that were 100-fold higher than wildtype SARS-CoV-2, suggesting that these variants may be more infectious.

A single study from China has reported that patient-derived variants of SARS-CoV-2 harboring mutations in the S glycoprotein demonstrate markedly different infectivity in vitro.[151] The authors used ultra-deep sequencing to characterize 11 viral isolates from a hospital near Wuhan, China. They identified 33 unique mutations (6 in the S glycoprotein), with some variants demonstrating significantly greater cytopathology in Vero E6 cells relative to other viral isolates. While this certainly demonstrates the potential for mutations to drastically alter SARS-CoV-2 pathogenicity, we should point out that these results are derived from an in vitro assay employing a non-human cell line. Furthermore, the studies assessing

the effects of S glycoprotein mutations on ACE2 binding affinity are currently limited to computational modeling.[149,150] The development of biological assays employing human cell-based systems are greatly needed in order to validate the results of these early studies and reliably assess the pathogenic potential of SARS-CoV-2 variants.

A recent report of a SARS-CoV-2 genome-wide SNP association study using 152 full length viral genomes identified a SNP at nucleotide 11 083 associated with COVID-19 severity. This SNP is located in the non-structural nsp6 protein.[152] The authors also identified three host miRNAs (miR-485-3p, miR-539-3p, and miR-3149) that target this region of the genome when it contains the 11083G variant but not when it contains 11083T; they speculate that the G variant may sequester these miRNA species, which leads to alterations in the biological pathways (cell proliferation and autophagy) controlled by these miRNAs.

7 | UNANSWERED QUESTIONS

Improving fundamental knowledge of SARS-CoV-2 and COVID-19 is a critical component of the NIAID Strategic Plan for COVID-19 Research initiative that is aligned with the priorities set by US Government to combat the burden of global COVID-19 pandemic.[153] Currently, we do not understand all of the contributing factors to the COVID-19 pandemic. Elucidating the underlying biology behind COVID-19 susceptibility is critical to mitigating this disease. Current evidence suggests that genetic variations contribute to inter-individual variability in immunity in humans. These human genetic factors (genetic variants) may affect innate and adaptive immune responses to SARS-CoV-2 and COVID-19 susceptibility. Genetic variants and their role in controlling immunity to SARS-CoV-2 remain unknown. Research efforts to gain a fundamental understanding of epigenetic mechanisms (eg, histone modifications, DNA methylation, and chromatin organization) in innate/cytokine immune response to SARS-CoV-2 infection and disease outcome are certainly needed and will help to inform the design of COVID-19 vaccine candidates.[154]

We now understand how two host cell factors, ACE2 and TMPRSS2, are being used by SARS-CoV-2 virus for cell entry and spike protein priming.[40] We are learning more about the concomitant proinflammatory cytokine "storm" phenomenon, particularly in young patients, resulting from COVID-19. A recent modeling study exploring structural analysis of human ACE2 polymorphic variants and SARS-CoV-2 S protein complexes has identified two ACE2 polymorphisms (rs73635825 [S19P] and rs143936283 [E329G]) that may have a functional impact on susceptibility/resistance against SARS-CoV-2 infection.[155] In this regard, further work is required to determine specific host genetic variants in *ACE2*, *TMPRSS2*, *CD147*, *MIF*, cytokine, and other SARS-CoV-2-related host genes and their functional effects that correspond with differential severity to COVID-19 infection (Figure 1). Consequently, innate, humoral, and cellular genetic risk signatures of infection and/or clinical disease severity can be developed. Replication studies will be critical to the evaluation of significant associations with specific variants in disease susceptibility genes.

The novel characteristics of SARS-CoV-2 and the general COVID-19 knowledge gap are impediments to the development of safe and effective vaccines against SARS-CoV-2, diagnostics/assays, antivirals, monoclonal antibodies, and outbreak control medical countermeasures that decrease diseases incidence and prevent mortality.[153] In particular, there is a critical need for both an improved understanding of protective immunity to SARS-CoV-2 and for effective and safe vaccines against COVID-19. Given the global spread of the COVID-19 pandemic, multiple vaccine candidates for SARS-CoV-2 are urgently being developed utilizing diverse technological platforms.[156,157] Understanding how SARS-CoV-2 interacts with host antiviral defense mechanisms must be considered a research priority, as new vaccines against COVID-19 are unlikely to be created without this fundamental knowledge. Furthermore, the immunological correlates of protection against COVID-19 remain undefined. Many questions must be answered concerning genetic variations, cellular components, and molecular mechanisms that are necessary for vaccine-induced protection and development of the memory CD4 + T, CD8 + T, and B cells. Measures of immune durability in different individuals to these new SARS-CoV-2 vaccines are critically needed. In the near future, we assert it will be feasible to identify interrelationships between humoral and cellular immunity and to identify a transcriptomic and/or functional molecular signature(s) of protective immunity to SARS-CoV-2.

Both host genetic differences and environmental factors may contribute to variability in immune responses to pathogens. For this reason, the outcome of COVID-19 is, in part, controlled by the SARS-CoV-2-host interaction. Clinical and serologic studies have demonstrated that SARS-CoV-2 infection and disease severity can be remarkably heterogeneous, with illness severity ranging from mild to critical.[158] The functional basis for this predisposition is currently unknown. It is essential to conduct population-based studies integrating immunogenetics with functional analysis of genetic variants to ensure the understanding of inter-individual variability in immune response to COVID-19 and candidate SARS-CoV-2 vaccines. Indeed, it is now clear that it is important to use the tools of molecular biology, genetics, immunology, bioinformatics, genetic epidemiology, and statistical genetics to comprehensively define how—and to what degree—inter-individual variations in SARS-CoV-2 susceptibility and subsequent immune responses are determined by gene polymorphisms.

To date, there are no systems biology studies that have examined the molecular risk signatures of SARS-CoV-2 infection and/or differential severity to COVID-19 infection, followed by a validation of the mechanism by which any identified signature functions. High-dimensional DNA (DNA-Seq) and RNA (RNA-Seq) sequencing may help discover mechanisms that control gene expression and gene/geneset molecular signatures associated with SARS-CoV-2 infection and disease severity. State-of-the-art "omics" technologies, such as single-cell mRNA sequencing, mass cytometry, proteomics, metabolomics, epigenetics, multiplex cytokine/chemokine measurements, bioinformatics, and others, will need to be applied to identify specific susceptibility markers and precise correlates of protection for immunity with applications to SARS-CoV-2 virus or vaccine-induced

immunity. Thus, current scientific efforts demonstrate a crucial need for COVID-19 host genetics studies.

8 | PROPOSED RESEARCH AGENDA

A critical step toward fully comprehending disease pathogenesis, prevention, and treatment will be attaining a comprehensive understanding of SARS-CoV-2 and human host genetics. Although it is an enormous and ongoing task, it is critical to understand viral evolution over time and geography, genetic epidemiology, immune response to virus and vaccines, and how and why different populations experience different manifestations and severity of disease. Such studies must be geographically diverse and take into consideration change over time. For example, as the SARS-CoV-2 virus accumulates mutations and unique recombination events, differences in clinical manifestations could result. Similarly, vaccines, monoclonal antibodies, and antivirals developed against one clade of the virus may be more or less effective against another clade of virus. Worse, it is possible that specific viral mutations could occur that increase the risk of antibody-enhanced disease in prior vaccinees over time.[159] Antibody-dependent enhancement (ADE) of disease occurs when non- or sub-neutralizing antibodies bind to a virus and promote its uptake into host cells via alternative mechanisms—most commonly through Fc receptors expressed on the surface of phagocytic immune cells. ADE has been proposed as a potential cause for the severity of COVID-19 in certain individuals, wherein mutations in the SARS-CoV-2 spike protein undermine the initial host antibody response and result in unstable virus-antibody complexes that promote infection of monocytes and macrophages in numerous tissues, leading to widespread apoptosis of immune cells and the development of a cytokine storm.[160] Indeed, acute lung injury was observed following viral challenge of vaccinated mice and non-human primates in trials of SARS-CoV-1 candidate vaccines,[161,162] and pronounced hepatitis was observed among vaccinated ferrets,[163] suggesting that ADE may be a concern for vaccine development efforts against SARS-CoV-2. A Th2-biased cellular immune response following vaccination can also predispose individuals to enhanced disease upon viral infection. The production of Th2-associated cytokines (eg, IL-4, IL-5, IL-13) potentiates granulocyte infiltration, enhanced mucus production, and dampened cytotoxic T cell activity in the respiratory tract, ultimately leading to impaired respiratory function and a significant reduction in viral clearance.[164,165] For these reasons, viral genetic sequencing studies over time and geography are critical.

Host genetic studies are much more expensive and complex. Such studies run from candidate gene to GWAS and epigenetic studies. Current observations suggest that although infection rates seem to be similar, disease severity appears to be worse in males vs females,[10,95-97,166] that racial differences in disease severity may also occur,[100,101,167] and that differences in morbidity and mortality vary by age group. The basis for these differences remains unclear.

From a viral and host genetics point of view, there are critical knowledge gaps that direct our research agenda. Examples include the following:

- Understanding the role of ACE2, TMPRSS2, and other specific candidate host genes on COVID-19 infection, severity, and disease outcome.
- Performing candidate gene, epigenetic, and GWAS studies across different racial and ethnic populations in order to identify genes and haplotypes associated with differential factors of infection and clinical outcome, as well as vaccine response.
- Systems genetic and biology studies to understand differential and interacting genetic effects on disease severity and outcome through the lens of the immune response network theory.[168]
- Studies of any differential effect of viral mutations and recombination events on differential disease severity in the context of host comorbidities and medication usage that blocks, suppresses, or activates differential host gene expression (eg, ACE2, TMPRSS2, etc).
- Understanding SAR-CoV-2 viral genetics over time and geography—particularly in terms of number and frequency of viral mutations and recombination events and their relationship to viral infectivity, transmissibility, disease severity, and clinical phenotype, viral burden, disease outcome, etc.
- Understanding potential interactions between SARS-CoV-2 and other respiratory co-infections (notably, between 20%-90% of COVID-19 patients in the United States were co-infected with at least one other respiratory pathogen).
- Understanding interactions and functional effects of various viral point mutations and recombination events on therapeutic monoclonal antibody, antiviral, and vaccine efficacy and possible adverse events.

ACKNOWLEDGEMENTS

We thank Caroline L. Vitse for her editorial assistance. This work was funded in parts by the National Institute of Allergy and Infectious Diseases grant numbers AI132348, AI121054, AI127365.

CONFLICT OF INTEREST

Dr Poland is the chair of a Safety Evaluation Committee for novel investigational vaccine trials being conducted by Merck Research Laboratories. Dr Poland offers consultative advice on vaccine development to Merck & Co. Inc, Avianax, Adjuvance, Valneva, Medicago, Sanofi Pasteur, GlaxoSmithKline, and Emergent Biosolutions. Drs. Poland and Ovsyannikova hold three patents related to measles and vaccinia peptide research. Dr Kennedy holds a patent on vaccinia peptide research. Dr Kennedy has received funding from Merck Research Laboratories to study waning immunity to measles and mumps after immunization with the MMR-II® vaccine. Drs. Poland, Kennedy, and Ovsyannikova have received grant funding from ICW Ventures for preclinical studies on a peptide-based COVID-19 vaccine. All other authors declare no competing financial interests. This research has been reviewed by the Mayo Clinic Conflict of Interest Review Board and was conducted in compliance with Mayo Clinic Conflict of Interest policies.

ORCID

Gregory A. Poland https://orcid.org/0000-0001-5057-4457

REFERENCES

1. Corman VM, Muth D, Niemeyer D, Drosten C. Hosts and sources of endemic human coronaviruses. *Adv Virus Res*. 2018;100:163-188.
2. Ogimi C, Kim YJ, Martin ET, Huh HJ, Chiu CH, Englund JA. What's new with the old coronaviruses? *J Pediatric Infect Dis Soc*. 2020;9(2):210-217.
3. Ogimi C, Greninger AL, Waghmare AA, et al. Prolonged shedding of human coronavirus in hematopoietic cell transplant recipients: risk factors and viral genome evolution. *J Infect Dis*. 2017;216(2):203-209.
4. de Wilde AH, Snijder EJ, Kikkert M, van Hemert MJ. Host factors in coronavirus replication. *Curr Top Microbiol Immunol*. 2018;419:1-42.
5. Drexler JF, Gloza-Rausch F, Glende J, et al. Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. *J Virol*. 2010;84(21):11336-11349.
6. Anthony SJ, Gilardi K, Menachery VD, et al. Further evidence for bats as the evolutionary source of middle east respiratory syndrome coronavirus. *mBio*. 2017;8(2):e00373-17.
7. Ahmed SF, Quadeer AA, McKay MR. Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV Immunological studies. *Viruses*. 2020;12(3):254.
8. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507-513.
9. Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA*. 2020;323:1061-1069.
10. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):497-506.
11. Novel Coronavirus Pneumonia Emergency Response Epidemiology T. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Zhonghua Liu Xing Bing Xue Za Zhi*. 2020;41(2):145-151.
12. CDC COVID-19 Response Team. Severe outcomes among patients with coronavirus disease 2019 (COVID-19) - United States, February 12-March 16, 2020. *MMWR*. 2020;69(12):343-346.
13. Han H, Yang L, Liu R, et al. Prominent changes in blood coagulation of patients with SARS-CoV-2 infection. *Clin Chem Lab Med*. 2020;58(7):1116-1120.
14. Kim IC, Kim JY, Kim HA, Han S. COVID-19-related myocarditis in a 21-year-old female patient. *Eur Heart J*. 2020;41(19):1859.
15. Li YC, Bai WZ, Hashikawa T. The neuroinvasive potential of SARS-CoV2 may play a role in the respiratory failure of COVID-19 patients. *J Med Virol*. 2020;92(6):552-555.
16. Conde Cardona G, Quintana Pajaro LD, Quintero Marzola ID, Ramos Villegas Y, Moscote Salazar LR. Neurotropism of SARS-CoV 2: mechanisms and manifestations. *J Neurol Sci*. 2020;412:116824.
17. Ottaviano G, Carecchio M, Scarpa B, Marchese-Ragona R. Olfactory and rhinological evaluations in SARS-CoV-2 patients complaining of olfactory loss. *Rhinology*. 2020. https://doi.org/10.4193/rhin20.136
18. Baig AM. Neurological manifestations in COVID-19 caused by SARS-CoV-2. *CNS Neurosci Ther*. 2020;26(5):499-501.
19. Xiao F, Tang M, Zheng X, Liu Y, Li X, Shan H. Evidence for Gastrointestinal Infection of SARS-CoV-2. *Gastroenterology*. 2020;158(6):1831-1833.e1833.
20. Liu R, Paxton WA, Choe S, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*. 1996;86(3):367-377.
21. Dendrou CA, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. *Nat Rev Immunol*. 2018;18(5):325-339.
22. Shen L, Wang C, Zhao J, et al. Delayed specific IgM antibody responses observed among COVID-19 patients with severe progression. *Emerg Microbes Infect*. 2020;9(1):1096-1101.
23. Sun B, Feng Y, Mo X, et al. Kinetics of SARS-CoV-2 specific IgM and IgG responses in COVID-19 patients. *Emerg Microbes Infect*. 2020;9(1):940-948.
24. Gralinski LE, Ferris MT, Aylor DL, et al. Genome wide identification of SARS-CoV susceptibility loci using the collaborative cross. *PLoS Genet*. 2015;11(10):e1005504.
25. Pyrc K, Berkhout B, van der Hoek L. The novel human coronaviruses NL63 and HKU1. *J Virol*. 2007;81(7):3051-3057.
26. Kahn JS. The widening scope of coronaviruses. *Curr Opin Pediatr*. 2006;18(1):42-47.
27. Berry M, Gamieldien J, Fielding BC. Identification of new respiratory viruses in the new millennium. *Viruses*. 2015;7(3):996-1019.
28. Ohtsuka N, Taguchi F. Mouse susceptibility to mouse hepatitis virus infection is linked to viral receptor genotype. *J Virol*. 1997;71(11):8860-8863.
29. Hirai A, Ohtsuka N, Ikeda T, et al. Role of mouse hepatitis virus (MHV) receptor murine CEACAM1 in the resistance of mice to MHV infection: studies of mice with chimeric mCEACAM1a and mCEACAM1b. *J Virol*. 2010;84(13):6654-6666.
30. Yeager CL, Ashmun RA, Williams RK, et al. Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature*. 1992;357(6377):420-422.
31. Wu K, Li W, Peng G, Li F. Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. *Proc Natl Acad Sci USA*. 2009;106(47):19970-19974.
32. Hofmann H, Simmons G, Rennekamp AJ, et al. Highly conserved regions within the spike proteins of human coronaviruses 229E and NL63 determine recognition of their respective cellular receptors. *J Virol*. 2006;80(17):8639-8652.
33. von Brunn A, Ciesek S, von Brunn B, Carbajo-Lozoya J. Genetic deficiency and polymorphisms of cyclophilin A reveal its essential role for Human Coronavirus 229E replication. *Curr Opin Virol*. 2015;14:56-61.
34. Luo C, Luo H, Zheng S, et al. Nucleocapsid protein of SARS coronavirus tightly binds to human cyclophilin A. *Biochem Biophys Res Commun*. 2004;321(3):557-565.
35. Neuman BW, Joseph JS, Saikatendu KS, et al. Proteomics analysis unravels the functional repertoire of coronavirus nonstructural protein 3. *J Virol*. 2008;82(11):5279-5294.
36. Li W, Moore MJ, Vasilieva N, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*. 2003;426(6965):450-454.
37. Li W, Sui J, Huang IC, et al. The S proteins of human coronavirus NL63 and severe acute respiratory syndrome coronavirus bind overlapping regions of ACE2. *Virology*. 2007;367(2):367-374.
38. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579(7798):270-273.
39. Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565-574.
40. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. 2020;181(2):271-280.
41. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*. 2020;181(2):281-292.
42. Chiu RW, Tang NL, Hui DS, et al. ACE2 gene polymorphisms do not affect outcome of severe acute respiratory syndrome. *Clin Chem*. 2004;50(9):1683-1686.

43. Itoyama S, Keicho N, Hijikata M, et al. Identification of an alternative 5'-untranslated exon and new polymorphisms of angiotensin-converting enzyme 2 gene: lack of association with SARS in the Vietnamese population. *Am J Med Genet A*. 2005;136(1):52-57.
44. Ansari MA, Marchi E, Ramamurthy N, et al. A gene locus that controls expression of ACE2 in virus infection. *medRxiv*. 2020. https://doi.org/10.1101/2020.04.26.20080408
45. Glowacka I, Bertram S, Muller MA, et al. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J Virol*. 2011;85(9):4122-4134.
46. Bertram S, Glowacka I, Muller MA, et al. Cleavage and activation of the severe acute respiratory syndrome coronavirus spike protein by human airway trypsin-like protease. *J Virol*. 2011;85(24):13363-13372.
47. Chan VS, Chan KY, Chen Y, et al. Homozygous L-SIGN (CLEC4M) plays a protective role in SARS coronavirus infection. *Nat Genet*. 2006;38(1):38-46.
48. Jeffers SA, Tusell SM, Gillim-Ross L, et al. CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. *Proc Natl Acad Sci USA*. 2004;101(44):15748-15753.
49. Pohlmann S, Zhang J, Baribaud F, et al. Hepatitis C virus glycoproteins interact with DC-SIGN and DC-SIGNR. *J Virol*. 2003;77(7):4070-4080.
50. Lin G, Simmons G, Pohlmann S, et al. Differential N-linked glycosylation of human immunodeficiency virus and Ebola virus envelope glycoproteins modulates interactions with DC-SIGN and DC-SIGNR. *J Virol*. 2003;77(2):1337-1346.
51. Pohlmann S, Soilleux EJ, Baribaud F, et al. DC-SIGNR, a DC-SIGN homologue expressed in endothelial cells, binds to human and simian immunodeficiency viruses and activates infection in trans. *Proc Natl Acad Sci USA*. 2001;98(5):2670-2675.
52. Marzi A, Gramberg T, Simmons G, et al. DC-SIGN and DC-SIGNR interact with the glycoprotein of Marburg virus and the S protein of severe acute respiratory syndrome coronavirus. *J Virol*. 2004;78(21):12090-12095.
53. Zhi L, Zhou G, Zhang H, et al. Lack of support for an association between CLEC4M homozygosity and protection against SARS coronavirus infection. *Nat Genet*. 2007;39(6):692-694; author reply 694–696.
54. Li H, Tang NL, Chan PK, et al. Polymorphisms in the C-type lectin genes cluster in chromosome 19 and predisposition to severe acute respiratory syndrome coronavirus (SARS-CoV) infection. *J Med Genet*. 2008;45(11):752-758.
55. Ip WK, Chan KH, Law HK, et al. Mannose-binding lectin in severe acute respiratory syndrome coronavirus infection. *J Infect Dis*. 2005;191(10):1697-1704.
56. Yuan FF, Tanner J, Chan PK, et al. Influence of FcgammaRIIA and MBL polymorphisms on severe acute respiratory syndrome. *Tissue Antigens*. 2005;66(4):291-296.
57. Zhang H, Zhou G, Zhi L, et al. Association between mannose-binding lectin gene polymorphisms and susceptibility to severe acute respiratory syndrome coronavirus infection. *J Infect Dis*. 2005;192(8):1355-1361.
58. Ruskamp JM, Hoekstra MO, Rovers MM, Schilder AG, Sanders EA. Mannose-binding lectin and upper respiratory tract infections in children and adolescents: a review. *Arch Otolaryngol Head Neck Surg*. 2006;132(5):482-486.
59. Tang F, Liu W, Zhang F, et al. IL-12 RB1 genetic variants contribute to human susceptibility to severe acute respiratory syndrome infection among Chinese. *PLoS One*. 2008;3(5):e2183.
60. Chan KY, Xu MS, Ching JC, et al. CD209 (DC-SIGN) -336A>G promoter polymorphism and severe acute respiratory syndrome in Hong Kong Chinese. *Hum Immunol*. 2010;71(7):702-707.
61. Chan KY, Ching JC, Xu MS, et al. Association of ICAM3 genetic variant with severe acute respiratory syndrome. *J Infect Dis*. 2007;196(2):271-280.
62. Chong WP, Ip WK, Tso GH, et al. The interferon gamma gene polymorphism +874 A/T is associated with severe acute respiratory syndrome. *BMC Infect Dis*. 2006;6:82.
63. Ng MW, Zhou G, Chong WP, et al. The association of RANTES polymorphism with severe acute respiratory syndrome in Hong Kong and Beijing Chinese. *BMC Infect Dis*. 2007;7:50.
64. Tu X, Chong WP, Zhai Y, et al. Functional polymorphisms of the CCL2 and MBL genes cumulatively increase susceptibility to severe acute respiratory syndrome coronavirus infection. *J Infect*. 2015;71(1):101-109.
65. Eisen DP. Mannose-binding lectin deficiency and respiratory tract infection. *J Innate Immun*. 2010;2(2):114-122.
66. Lau YL, Peiris JS. Pathogenesis of severe acute respiratory syndrome. *Curr Opin Immunol*. 2005;17(4):404-410.
67. Eisen DP, Dean MM, Boermeester MA, et al. Low serum mannose-binding lectin level increases the risk of death due to pneumococcal infection. *Clin Infect Dis*. 2008;47(4):510-516.
68. Roy S, Knox K, Segal S, et al. MBL genotype and risk of invasive pneumococcal disease: a case-control study. *Lancet*. 2002;359(9317):1569-1573.
69. Areeshi MY, Mandal RK, Akhter N, et al. A meta-analysis of MBL2 polymorphisms and tuberculosis risk. *Sci Rep*. 2016;6:35728.
70. Faber J, Schuessler T, Finn A, et al. Age-dependent association of human mannose-binding lectin mutations with susceptibility to invasive meningococcal disease in childhood. *Pediatr Infect Dis J*. 2007;26(3):243-246.
71. Lau YL, Peiris JS. Association of cytokine and chemokine gene polymorphisms with severe acute respiratory syndrome. *Hong Kong Med J*. 2009;15(Suppl 2):43-46.
72. Smith CA, Tyrell DJ, Kulkarni UA, et al. Macrophage migration inhibitory factor enhances influenza-associated mortality in mice. *JCI Insight*. 2019;4(13). e128034.
73. Yende S, Angus DC, Kong L, et al. The influence of macrophage migration inhibitory factor gene polymorphisms on outcome from community-acquired pneumonia. *FASEB J*. 2009;23(8):2403-2411.
74. Savva A, Brouwer MC, Roger T, et al. Functional polymorphisms of macrophage migration inhibitory factor as predictors of morbidity and mortality of pneumococcal meningitis. *Proc Natl Acad Sci USA*. 2016;113(13):3597-3602.
75. Lin M, Tseng HK, Trejaut JA, et al. Association of HLA class I with severe acute respiratory syndrome coronavirus infection. *BMC Med Genet*. 2003;4:9.
76. Ng MH, Lau KM, Li L, et al. Association of human-leukocyte-antigen class I (B*0703) and class II (DRB1*0301) genotypes with susceptibility and resistance to the development of severe acute respiratory syndrome. *J Infect Dis*. 2004;190(3):515-518.
77. Chen YM, Liang SY, Shih YP, et al. Epidemiological and genetic correlates of severe acute respiratory syndrome coronavirus infection in the hospital with the highest nosocomial infection rate in Taiwan in 2003. *J Clin Microbiol*. 2006;44(2):359-365.
78. Nguyen A, David JK, Maden SK, et al. Human leukocyte antigen susceptibility map for SARS-CoV-2. *J Virol*. 2020;94(13), e00510–20.
79. Ng MH, Cheng SH, Lau KM, et al. Immunogenetics in SARS: a case-control study. *Hong Kong Med J*. 2010;16(5 Suppl 4):29-33.
80. Keicho N, Itoyama S, Kashiwase K, et al. Association of human leukocyte antigen class II alleles with severe acute respiratory syndrome in the Vietnamese population. *Hum Immunol*. 2009;70(7):527-531.
81. Wang SF, Chen KH, Chen M, et al. Human-leukocyte antigen class I Cw 1502 and class II DR 0301 genotypes are associated with resistance to severe acute respiratory syndrome (SARS) infection. *Viral Immunol*. 2011;24(5):421-426.

82. Li X, Geng M, Peng Y, Meng L, Lu S. Molecular immune pathogenesis and diagnosis of COVID-19. *J Pharm Anal*. 2020;10(2):102-108.
83. Yuan FF, Velickovic Z, Ashton LJ, et al. Influence of HLA gene polymorphisms on susceptibility and outcome post infection with the SARS-CoV virus. *Virol Sin*. 2014;29(2):128-130.
84. Xiong P, Zeng X, Song MS, et al. Lack of association between HLA-A, -B and -DRB1 alleles and the development of SARS: a cohort of 95 SARS-recovered individuals in a population of Guangdong, southern China. *Int J Immunogenet*. 2008;35(1):69-74.
85. Widagdo W, Sooksawasdi Na Ayudhya S, Hundie GB, Haagmans BL. Host determinants of MERS-CoV transmission and pathogenesis. *Viruses*. 2019;11(3):280.
86. Widagdo W, Okba NMA, Li W, et al. Species-specific colocalization of middle east respiratory syndrome coronavirus attachment and entry receptors. *J Virol*. 2019;93(16).e00107–19.
87. van Doremalen N, Miazgowicz KL, Milne-Price S, et al. Host species restriction of Middle East respiratory syndrome coronavirus through its receptor, dipeptidyl peptidase 4. *J Virol*. 2014;88(16):9220-9232.
88. Hajeer AH, Balkhy H, Johani S, Yousef MZ, Arabi Y. Association of human leukocyte antigen class II alleles with severe Middle East respiratory syndrome-coronavirus infection. *Ann Thorac Med*. 2016;11(3):211-213.
89. Van Dyke AL, Cote ML, Wenzlaff AS, Land S, Schwartz AG. Cytokine SNPs: Comparison of allele frequencies by race and implications for future studies. *Cytokine*. 2009;46(2):236-244.
90. Maiers M, Gragert L, Klitz W. High-resolution HLA alleles and haplotypes in the United States population. *Hum Immunol*. 2007;68(9):779-788.
91. Buhler S, Sanchez-Mazas A. HLA DNA sequence variation among human populations: molecular signatures of demographic and selective events. *PLoS One*. 2011;6(2):e14643.
92. de Magalhaes JP, Curado J, Church GM. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics*. 2009;25(7):875-881.
93. Grath S, Parsch J. Sex-biased gene expression. *Annu Rev Genet*. 2016;50:29-44.
94. Ellegren H, Parsch J. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet*. 2007;8(9):689-698.
95. Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. 2020;382(18):1708-1720.
96. Conti P, Younes A. Coronavirus COV-19/SARS-CoV-2 affects women less than men: clinical response to viral infection. *J Biol Regul Homeost Agents*. 2020;34(2). https://doi.org/10.23812/Editorial-Conti-3
97. Gudbjartsson DF, Helgason A, Jonsson H, et al. Spread of SARS-CoV-2 in the icelandic population. *N Engl J Med*. 2020;382(24):2302-2315.
98. Zhao J, Yang Y, Huang H, et al. Relationship between the ABO blood group and the COVID-19 susceptibility. *medRxiv*. 2020. https://doi.org/10.1101/2020.03.11.20031096
99. Zietz M, Tatonetti NP. Testing the association between blood type and COVID-19 infection, intubation, and death. *medRxiv*. 2020. https://doi.org/10.1101/2020.04.08.20058073
100. Garg S, Kim L, Whitaker M, et al. Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019 - COVID-NET, 14 States, March 1–30, 2020. *MMWR*. 2020;69(15):458-464.
101. Gold JAW, Wong KK, Szablewski CM, et al. Characteristics and clinical outcomes of adult patients hospitalized with COVID-19 - Georgia, March 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(18):545-550.
102. Zeng F, Dai C, Cai P, et al. A comparison study of SARS-CoV-2 IgG antibody between male and female COVID-19 patients: a possible reason underlying different outcome between sex. *J Med Virol*. 2020. https://doi.org/10.1002/jmv.25989
103. Benetti E, Tita R, Spiga O, et al. ACE2 variants underlie interindividual variability and susceptibility to COVID-19 in Italian population. *medRxiv*. 2020. https://doi.org/10.1101/2020.04.03.20047977
104. Stawiski EW, Diwanji D, Suryamohan K, et al. Human ACE2 receptor polymorphisms predict SARS-CoV-2 susceptibility. *bioRxiv*. 2020. https://doi.org/10.1101/2020.04.07.024752
105. Li Q, Cao Z, Rahman P. Genetic variability of human angiotensin-converting enzyme 2 (hACE2) among various ethnic populations. *bioRxiv*. 2020. https://doi.org/10.1101/2020.04.14.041434
106. Pinto BG, Oliveira AE, Singh Y, et al. ACE2 expression is increased in the lungs of patients with comorbidities associated with severe COVID-19. *medRxiv*. 2020. https://doi.org/10.1101/2020.03.21.20040261
107. Cristiani L, Mancino E, Matera L, et al. Will children reveal their secret? The coronavirus dilemma. *Eur Respir J*. 2020;55(4):2000749.
108. Xie X, Chen J, Wang X, Zhang F, Liu Y. Age- and gender-related difference of ACE2 expression in rat lung. *Life Sci*. 2006;78(19):2166-2171.
109. Kuba K, Imai Y, Rao S, Jiang C, Penninger JM. Lessons from SARS: control of acute lung failure by the SARS receptor ACE2. *J Mol Med*. 2006;84(10):814-820.
110. Tukiainen T, Villani AC, Yen A, et al. Landscape of X chromosome inactivation across human tissues. *Nature*. 2017;550(7675):244-248.
111. Zhao Y, Zhao Z, Wang Y, Zhou Y, Ma Y, Zuo W. Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov. *bioRxiv*. 2020.
112. Asselta R, Paraboschi EM, Mantovani A, Duga S. ACE2 and TMPRSS2 variants and expression as candidates to sex and country differences in COVID-19 severity in Italy. *medRxiv*. 2020. https://doi.org/10.1101/2020.03.30.20047878
113. Cai G. Bulk and single-cell transcriptomics identify tobacco-use disparity in lung gene expression of ACE2, the receptor of 2019-nCov. *medRxiv*. 2020. https://doi.org/10.1101/2020.02.05.20020107
114. Gibson WT, Evans DM, An J, Jones SJ. ACE 2 coding variants: a potential X-linked risk factor for COVID-19 disease. *bioRxiv*. 2020. https://doi.org/10.1101/2020.04.05.026633
115. Iyer SP, Ensor J, Anand K, et al. Higher mortality in men from COVID19 infection-understanding the factors that drive the differences between the biological sexes. *medRxiv*. 2020. https://doi.org/10.1101/2020.04.19.20062174
116. Gebhard C, Regitz-Zagrosek V, Neuhauser HK, Morgan R, Klein SL. Impact of sex and gender on COVID-19 outcomes in Europe. *Biol Sex Differ*. 2020;11(1):29.
117. Russo R, Andolfo I, Lasorsa VA, Iolascon A, Capasso M. Genetic analysis of the novel SARS-CoV-2 host receptor *TMPRSS2* in different populations. *bioRxiv*. 2020. https://doi.org/10.1101/2020.04.23.057190
118. Warren RL, Birol I. HLA predictions from the bronchoalveolar lavage fluid samples of five patients at the early stage of the Wuhan seafood market COVID-19 outbreak. *arXiv*. 2004.07108 [q-bio.GN]. 2020.
119. Hyun-Jung Lee C, Koohy H. In silico identification of vaccine targets for 2019-nCoV. *F1000Res*. 2020;9:145.
120. Campbell KM, Steiner G, Wells DK, Ribas A, Kalbasi A. Prediction of SARS-CoV-2 epitopes across 9360 HLA class I alleles. *bioRxiv*. 2020. https://doi.org/10.1101/2020.03.30.016931
121. Ellinghaus D, Degenhardt F, Bujanda L, et al. The ABO blood group locus and a chromosome 3 gene cluster associate with SARS-CoV-2 respiratory failure in an Italian-Spanish genome-wide association analysis. *medRxiv*. 2020. https://doi.org/10.1101/2020.05.31.20114991
122. Gubernatorova EO, Gorshkova EA, Polinova AI, Drutskaya MS. IL-6: relevance for immunopathology of SARS-CoV-2. *Cytokine Growth Factor Rev*. 2020;6:13-24.
123. Kadkhoda K. COVID-19: an immunopathological view. *mSphere*. 2020;5(2).e00344–20.

124. Chen G, Wu D, Guo W, et al. Clinical and immunologic features in severe and moderate coronavirus disease 2019. *J Clin Invest*. 2020;130(5):2620-2629.
125. Hung IF, Lung KC, Tso EY, et al. Triple combination of interferon beta-1b, lopinavir-ritonavir, and ribavirin in the treatment of patients admitted to hospital with COVID-19: an open-label, randomised, phase 2 trial. *Lancet*. 2020;395(10238):1695-1704.
126. Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Med*. 2020;46(5):846-848.
127. Herold T, Jurinovic V, Arnreich C, et al. Elevated levels of IL-6 and CRP predict the need for mechanical ventilation in COVID-19. *J Allergy Clin Immunol*. 2020.S0091-6749(20), 30685.
128. Ulhaq ZS, Soraya GV. Interleukin-6 as a potential biomarker of COVID-19 progression. *Med Mal Infect*. 2020;50(4):382-383.
129. Ulhaq ZS, Soraya GV. Anti-IL-6 Receptor Antibody Treatment for Severe COVID-19 and the Potential Implication of IL-6 Gene Polymorphisms in Novel Coronavirus Pneumonia (May 2, 2020). SSRN: https://ssrn.com/abstract=3592878 or 2020, 10.2139/ssrn.3592878.
130. Kirtipal N, Bharadwaj S. Interleukin 6 polymorphisms as an indicator of COVID-19 severity in humans. *J Biomol Struct Dyn*. 2020;1-4. https://doi.org/10.1080/07391102.2020.1776640
131. The COVID-19 host genetics initiative. https://www.covid19hg.org/ 2020.
132. COVID-19 Host Genetics Initiative. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eur J Hum Genet*. 2020;28(6):715-718.
133. COVID human genetic effort. https://www.covidhge.com 2020.
134. National Institute for Health Research. Genetics of susceptibility and mortality in critical care (GenOMICC). https://www.nihr.ac.uk/covid-studies/study-detail.htm?entryId=189676 2020.
135. Sanjuan R, Domingo-Calap P. Mechanisms of viral mutation. *Cell Mol Life Sci*. 2016;73(23):4433-4448.
136. Narasimhan VM, Rahbari R, Scally A, et al. Estimating the human mutation rate from autozygous segments reveals population differences in human mutational processes. *Nat Commun*. 2017;8(1):303.
137. Karamitros T, Papadopoulou G, Bousali M, Mexias A, Tsiodras S, Mentis A. SARS-CoV-2 exhibits intra-host genomic plasticity and low-frequency polymorphic quasispecies. *bioRxiv*. 2020. https://doi.org/10.1101/2020.03.27.009480
138. Karlsson EK, Kwiatkowski DP, Sabeti PC. Natural selection and infectious disease in human populations. *Nat Rev Genet*. 2014;15(6):379-393.
139. Forster P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc Natl Acad Sci U S A*. 2020;117(17):9241-9243.
140. The Nextstrain Team. Genomic epidemiology of novel coronavirus-Global sampling. https://nextstrain.org/ncov/global 2020.
141. Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*. 2018;34(23):4121-4123.
142. Schleunes A. Relatively Stable SARS-CoV-2 Genome Is Good News for a Vaccine. The Scientist 2020. https://www.the-scientist.com/news-opinion/relatively-stable-sars-cov-2-genome-is-good-news-for-a-vaccine-67319
143. Genomics Research Technology. https://www.technologynetworks.com/genomics/product-news/sars-cov-2-gene-variants-provide-clues-to-coronavirus-epidemiology-332573 March 25, 2020.
144. Korber B, Fischer W, Gnanakaran S, et al. Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. *bioRxiv*. 2020. https://doi.org/10.1101/2020.04.29.069054
145. Becerra-Flores M, Cardozo T. SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate. *Int J Clin Pract*. 2020. https://doi.org/10.1111/ijcp.13525
146. Pachetti M, Marini B, Benedetti F, et al. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J Transl Med*. 2020;18(1):179.
147. Agostini ML, Andres EL, Sims AC, et al. Coronavirus susceptibility to the antiviral remdesivir (GS-5734) is mediated by the viral polymerase and the proofreading exoribonuclease. *mBio*. 2018;9(2).e00221–18.
148. Goldhill DH, Te Velthuis AJW, Fletcher RA, et al. The mechanism of resistance to favipiravir in influenza. *Proc Natl Acad Sci USA*. 2018;115(45):11613-11618.
149. Jia Y, Shen G, Zhang Y, et al. Analysis of the mutation dynamics of SARS-CoV-2 reveals the spread history and emergence of RBD mutant with lower ACE2 binding affinity. *bioRxiv*. 2020. https://doi.org/10.1101/2020.04.09.034942
150. Ou J, Zhou Z, Dai R, et al. Emergence of RBD mutations in circulating SARS-CoV-2 strains enhancing the structural stability and human ACE2 receptor affinity of the spike protein. *bioRxiv*. 2020. https://doi.org/10.1101/2020.03.15.991844
151. Yao H, Lu X, Chen Q, et al. Patient-derived mutations impact pathogenicity of SARS-CoV-2. *medRxiv*. 2020. https://doi.org/10.1101/2020.04.14.20060160
152. Aiewsakun P, Wongtrakoongate P, Thawornwattana Y, Hongeng S, Thitithanyanont A. SARS-CoV-2 genetic variations associated with COVID-19 severity. *medRxiv*. 2020. https://doi.org/10.1101/2020.05.27.20114546
153. National Institute of Allergy and Infectious Diseases. NIAID strategic plan for COVID-19 research. April 22, 2020. https://www.niaid.nih.gov/sites/default/files/NIAID-COVID-19-Strategic-Plan-2020.pdf.2020
154. Schafer A, Baric RS. Epigenetic landscape during coronavirus infection. *Pathogens*. 2017;6:1.
155. Hussain M, Jabeen N, Raza F, et al. Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. *J Med Virol*. 2020. https://doi.org/10.1002/jmv.25832
156. Le Thanh T, Andreadakis Z, Kumar A, et al. The COVID-19 vaccine development landscape. *Nat Rev Drug Discov*. 2020;19(5):305-306.
157. Lurie N, Saville M, Hatchett R, Halton J. Developing Covid-19 vaccines at pandemic speed. *N Engl J Med*. 2020;382(21):1969-1973.
158. Wu Z, McGoogan JM. Characteristics of and Important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the chinese center for disease control and prevention. *JAMA*. 2020;323(13):1239.
159. Poland GA. Tortoises, hares, and vaccines: a cautionary note for SARS-CoV-2 vaccine development. *Vaccine*. 2020;38(27):4219-4220.
160. Ricke D, Malone RW. Medical countermeasures analysis of 2019-nCoV and vaccine risks for antibody-dependent enhancement (ADE) (2/27/2020). Preprint available at SSRN: https://ssrn.com/abstract=3546070 or https://doi.org/10.2139/ssrn.3546070 In:2020.
161. Tseng CT, Sbrana E, Iwata-Yoshikawa N, et al. Immunization with SARS coronavirus vaccines leads to pulmonary immunopathology on challenge with the SARS virus. *PLoS One*. 2012;7(4):e35421.
162. Liu L, Wei Q, Lin Q, et al. Anti-spike IgG causes severe acute lung injury by skewing macrophage responses during acute SARS-CoV infection. *JCI Insight*. 2019;4(4).e123158
163. Weingartl H, Czub M, Czub S, et al. Immunization with modified vaccinia virus Ankara-based recombinant vaccine against severe acute respiratory syndrome is associated with enhanced hepatitis in ferrets. *J Virol*. 2004;78(22):12672-12676.
164. Graham BS. Rapid COVID-19 vaccine development. *Science*. 2020;368(6494):945-946.

165. Ruckwardt TJ, Morabito KM, Graham BS. Immunological lessons from respiratory syncytial virus vaccine development. *Immunity*. 2019;51(3):429-442.
166. Jin JM, Bai P, He W, et al. Gender differences in patients with COVID-19: Focus on severity and mortality. *Front Public Health*. 2020;8:152.
167. McCoy J, Wambier CG, Vano-Galvan S, et al. Racial variations in COVID-19 deaths may be due to androgen receptor genetic variants associated with prostate cancer and androgenetic alopecia. Are anti-androgens a potential treatment for COVID-19? *J Cosmet Dermatol*. 2020;19(7):1542-1543.
168. Poland GA, Kennedy RB, McKinney BA, et al. Vaccinomics, adversomics, and the immune response network theory: individualized vaccinology in the 21st century. *Semin Immunol*. 2013;25(2):89-103.

How to cite this article: Ovsyannikova IG, Haralambieva IH, Crooke SN, Poland GA, Kennedy RB. The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity. *Immunol Rev*. 2020;296:205–219. https://doi.org/10.1111/imr.12897